

B. Loeb

Re-run

#8  
mt  
10/9/01

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/668,119

DATE: 10/09/2001  
TIME: 16:53:14

Input Set : A:\01100250.app  
Output Set: N:\CRF3\10092001\I668119.raw

C--> 3 <110> APPLICANT: Solomon, William B  
4 Abraham, Shaji  
6 <120> TITLE OF INVENTION: Transcriptional adaptor protein  
8 <130> FILE REFERENCE: 011.00250  
10 <140> CURRENT APPLICATION NUMBER: US 09/668,119  
11 <141> CURRENT FILING DATE: 2001-09-10  
13 <160> NUMBER OF SEQ ID NOS: 8  
15 <170> SOFTWARE: PatentIn Ver. 2.1  
17 <210> SEQ ID NO: 1  
18 <211> LENGTH: 1740  
19 <212> TYPE: DNA  
20 <213> ORGANISM: Homo sapiens  
22 <400> SEQUENCE: 1  
23 atgaggaaaag ctggtgtggc acacagtaaa tccagcaagg atatggagag ccatgttttc 60  
24 ctgaaggcca agaccggga cgaatacctt tctctcgtgg ccaggctcat tatccatttt 120  
25 cgagacattc ataacaagaa atctcaagct tccgtcagtg atcctatgaa tgcactccag 180  
26 agcctgactg gcggacctgc tgcgggagcc gctggaattg gcatgcctcc tcggggcccg 240  
27 ggacagtctc tgggcgggat gggtagcttt ggtgccatgg gacagccaat gtctctctca 300  
28 gggcagccgc ctctggggac ctcggggatg gcccctcaca gcatggctgt cgtgtctacg 360  
29 gcaactccac agaccagct gcagctccag cagggtggcg tgcagcagca gcagcaacag 420  
30 cagcagttcc agcagcagca gcaggcggcg ctacagcagc agcagcagca gcagcaacag 480  
31 cagcagttcc aggtcagca gagtgcctat cagcagcagt tccaagcagt agtgcagcag 540  
32 cagcagcagc tccagcagca gcagcagcag cagcagcatc taattaaatt gcatcatcaa 600  
33 aatcagcaac agatacagca gcagcaacag cagctgcagc gaatagcaca gctgcagctc 660  
34 caacaacagc aacagcagca gcagcagcag cagcagcagc agcagcagc tttggaggcc 720  
35 cagccaccaaa ttcagcagcc accgatgcag cagccacagc ctccgccctc ccaggctctg 780  
36 ccccagcagc tgcagcagat gcatcacaca cagcaccacc agccgccacc acagccccag 840  
37 cagcctccag ttgtcagaa ccaaccatca caactcccgc cacagtgcga gaccagcct 900  
38 ttggtgtcac aggcgcaagc tctccctgga caaatgttgt ataccaacc accactgaaa 960  
39 tttgtccgag ctccgatggg ggtgcagcag ccccagtgcc agccccaggt gcagcagcag 1020  
40 cagacagcag tacagacagc tcaggctgcc cagatgggtg ctcccggagt ccaggtcagc 1080  
41 cagagcagcc tccccatgct gtccctgccg tcaccgggcc agcaggtgca gaccccgagc 1140  
42 tcgatgcccc ctcccccca gccgtccccg cagcccggcc agcccagctc acagcccaac 1200  
43 tccaacgtca gctctggccc tgccccatct cccagtagct tctgcccag cccctcaccg 1260  
44 cagccctccc agagcccagt gacggcgagg accccacaga acttcagtgt cccctcacct 1320  
45 ggacctttta acacacctgt gaaccccagc tctgtcatga gccagctgg ctccagccag 1380  
46 gctgaggagc agcagtacct ggacaagctg aagcagctgt cgaagtacat cgagccccctg 1440  
47 cgccgcatga tcaacaagat cgacaagaa gaagacagaa aaaaggacct gagtaagatg 1500  
48 aagagccttc tggacattct gacagacccc tcgaagcggg gtccccctgaa gaccttgcaa 1560  
49 aagtgtgaga tgcctctgga gaaactcaag aatgacatgc ggtgcccact cccccaccgc 1620  
50 ccccggtgcc accgacaaa cagcagttac tatgccagcc gctcctggat gccgtcctgg 1680  
51 ccaacatccg ctacactgtc ttcaaccatt cctgtaccg cacattcgtt ccagccatga 1740  
54 <210> SEQ ID NO: 2  
55 <211> LENGTH: 3334  
56 <212> TYPE: DNA  
57 <213> ORGANISM: Homo sapiens  
59 <400> SEQUENCE: 2

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/668,119

DATE: 10/09/2001  
TIME: 16:53:14

Input Set : A:\01100250.app  
Output Set : N:\CRF3\10092001\I668119.raw

```

60 ccggcgcgac ttgggcctgg ctctgtgact gaggcgggcg cggtggcggc caagcgggat 60
61 acggggcgcg ggagctgggg aacaggcatg gacgtttccg ggcaagagac cgactggcgg 120
62 agaccgcctt ccggcagaag ctggtcagtc aaatcgagga tgccatgagg aaagctggtg 180
63 tggcacacag taaatccagc aaggatatgg agagccatgt tttcctgaag gccaagacct 240
64 gggacgaata cttttctctc gtggccaggc tcattatcca ttttcgagac attcataaca 300
65 agaaatctca agcttccgtc agtgatccta tgaatgcaact ccagagcctg actggcgagc 360
66 ctgctgcggg agccgctgga attggcatgc ctccctcgggg cccgggacag tctctgggcg 420
67 ggatgggtag ctttggtgcc atgggacagc caatgtctct ctacgggagc ccgcctcctg 480
68 ggacctcggg gatggccctt cacagcatgg agtcgtgtc tacggcaact ccacagacct 540
69 agctgcagct ccagcagggt gcgctgcagc agcagcagca acagcagcag ttccagcagc 600
70 agcagcaggc ggcgtacag cagcagcagc agcagcagca acagcagcag ttccaggctc 660
71 agcagagtgc catgcagcag cagttccaag cagtagtgca gcagcagcag cagctccagc 720
72 agcagcagca gcagcagcag catctaatta aattgcatca tcaaatcag caacagatac 780
73 agcagcagca acagcagctg cagcgaatag cacagctgca gctccaacaa cagcaacagc 840
74 agcagcagca gcagcagcag cagcagcagc aggttttggg ggcccagcca ccaattcagc 900
75 agccaccgat gcagcagcca cagcctccgc cctcccaggc tctgccccag cagctgcagc 960
76 agatgcacac cacacagcac caccagccgc caccacagcc ccagcagcct ccagttgctc 1020
77 agaaccaacc atcacaaact ccgccacagt cgcagaccca gcctttggtg tcacaggcgc 1080
78 aagctctccc tggacaaatg ttgtataccc aaccaccact gaaatttgc cgagctccga 1140
79 tgggtggtgca gcagccccc gtgcagcccc aggtgcagca gcagcagaca gcagtagaca 1200
80 cagctcaggc tgcccagatg gtggtccccg gagtccagggt cagccagagc agcctcccca 1260
81 tgctgtcttc gcgctcaccg ggccagcagg tgcagacccc gcagtcgatg cccctcccc 1320
82 cccagccgct cccgcagccc ggccagccca gctcacagcc caactccaac gtcagctctg 1380
83 gccctgcccc atctcccagt agcttccctg ccagccccc accgcagccc tcccagagcc 1440
84 cagtgcaggc ggggacccca cagaacttca gtgtccccc acctggacct ttaaacacac 1500
85 ctgtgaaccc cagctctgtc atgagcccag ctggtccag ccaggctgag gagcagcagt 1560
86 acctggacaa gctgaagcag ctgtcgaagt acatcgagcc cctgcgccc atgatcaaca 1620
87 agatcgacaa gaacgaagac agaaaaaagg acctgagtaa gatgaagagc cttctggaca 1680
88 ttctgacaga cccctcgaag cgggtgtccc tgaagacctt gcaaaaagtgt gagatcgccc 1740
89 tggagaaact caagaatgac atgcggtgcc cactccccca ccgcccccg tgccaccgac 1800
90 caaacagcag tacctatgcc agccgctcct ggatgccgtc ctggccaaca tccgctcacc 1860
91 tgtcttcaac cattccctgt accgcacatt cgttccagcc atgaccgcca ttcacggccc 1920
92 acccatcacg gcccagtggt tgtgcacccg gaagcgcagg cttgaggatg atgagcgcca 1980
93 gagcatcccc agtgtgtctc aggggtgaggt ggccaggctg gaccccaagt tccgtgtaaa 2040
94 cctggacctt tctactgca gcaacaatgg cactgtccac ctgatctgca agctggatga 2100
95 caaggacctc ccaagtgtgc caccactgga gtcagtgtg cccgctgact atcctgccc 2160
96 aagcccgtct tggatagacc ggcagtggca gtacgacgcc aaccccttct ctcagctcg 2220
97 tgcaccgtct catgacctcc aggtctgtgc agctcccga caagcactcg gtcaccgct 2280
98 tgcctaacac ctgggcccag agcgtccacc aggcctgct ctcagccgc tagccaagac 2340
99 tgcagggatg gcccgagcc tcatcgggc caaggacaca cgctcctgt cagacacttc 2400
100 taggtgttg cttccttaga gagcctgggg ttaggttagc tttcctgctt ttatcttctg 2460
101 ccttggggac ctgcccacg aaatcccaca cctgtacaga actgggatag gcgcagtgga 2520
102 gggggttgct tggggggcgt tggccgactt cttagagaag gccctccatg tgacttctc 2580
103 ccaggagcca gatgcgatcc tcaggctgct ctcaccgtgg cctgtccacg gtccaggctc 2640
104 atctcagcag cgtgaggggt cactcagggt gttgttagag cgtctcgtgt gtgctagacg 2700
105 caccctact cgttctctata gaacacagag gacataggaa acccttaaaa cacacatggg 2760
106 attctctggt cacagttttg ggttcaggct atgctgcttt gggcagggtg agcaccctcc 2820
107 gaggaagcct gcaagtccag ggcacaggct gccttttggg gggagggctg gcccataggt 2880
108 gctgctggt ccccgccacc agctgggcct cagccctcac ggcattcctg ctgagcaccg 2940

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/668,119

DATE: 10/09/2001  
TIME: 16:53:14

Input Set : A:\01100250.app  
Output Set: N:\CRF3\10092001\I668119.raw

```

109 tggggcacc agggagcagg ggcgtcaggg atcctgctgc cggcaccct gtgccgctgg 3000
110 catgagggcc gtgtcccccac tgtgaaggat gaagagcaag gccctcagga cccgtgtcct 3060
111 cagagcacca cacactgagc acccagagac agcgggcttg gcagcgggcc gggccatgca 3120
112 gggagcgcct ccctatgttg cctgccactc tgggcaccgg ccagcaccct ctggtgagaa 3180
113 gaggtccccc ctttttatgt gcactacccc accatctgtg attataataa atttattatt 3240
114 cctgtggaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3300
115 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 3334
118 <210> SEQ ID NO: 3
119 <211> LENGTH: 579
120 <212> TYPE: PRT
121 <213> ORGANISM: Homo sapiens
123 <400> SEQUENCE: 3
124 Met Arg Lys Ala Gly Val Ala His Ser Lys Ser Ser Lys Asp Met Glu
125 1 5 10 15
127 Ser His Val Phe Leu Lys Ala Lys Thr Arg Asp Glu Tyr Leu Ser Leu
128 20 25 30
130 Val Ala Arg Leu Ile Ile His Phe Arg Asp Ile His Asn Lys Lys Ser
131 35 40 45
133 Gln Ala Ser Val Ser Asp Pro Met Asn Ala Leu Gln Ser Leu Thr Gly
134 50 55 60
136 Gly Pro Ala Ala Gly Ala Ala Gly Ile Gly Met Pro Pro Arg Gly Pro
137 65 70 75 80
139 Gly Gln Ser Leu Gly Gly Met Gly Ser Phe Gly Ala Met Gly Gln Pro
140 85 90 95
142 Met Ser Leu Ser Gly Gln Pro Pro Pro Gly Thr Ser Gly Met Ala Pro
143 100 105 110
145 His Ser Met Ala Val Val Ser Thr Ala Thr Pro Gln Thr Gln Leu Gln
146 115 120 125
148 Leu Gln Gln Val Ala Leu Gln Gln Gln Gln Gln Gln Gln Phe Gln
149 130 135 140
151 Gln Gln Gln Ala Ala Leu Gln Gln Gln Gln Gln Gln Gln Gln
152 145 150 155 160
154 Gln Gln Phe Gln Ala Gln Gln Ser Ala Met Gln Gln Gln Phe Gln Ala
155 165 170 175
157 Val Val Gln Gln Gln Gln Gln Leu Gln Gln Gln Gln Gln Gln Gln
158 180 185 190
160 His Leu Ile Lys Leu His His Gln Asn Gln Gln Gln Ile Gln Gln Gln
161 195 200 205
163 Gln Gln Gln Leu Gln Arg Ile Ala Gln Leu Gln Gln Gln Gln Gln
164 210 215 220
166 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Ala Leu Glu Ala
167 225 230 235 240
169 Gln Pro Pro Ile Gln Gln Pro Pro Met Gln Gln Pro Gln Pro Pro Pro
170 245 250 255
172 Ser Gln Ala Leu Pro Gln Gln Leu Gln Gln Met His His Thr Gln His
173 260 265 270
175 His Gln Pro Pro Pro Gln Pro Gln Gln Pro Pro Val Ala Gln Asn Gln
176 275 280 285
178 Pro Ser Gln Leu Pro Pro Gln Ser Gln Thr Gln Pro Leu Val Ser Gln

```

10/9/01

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/668,119

DATE: 10/09/2001  
TIME: 16:53:14

Input Set : A:\01100250.app  
Output Set: N:\CRF3\10092001\I668119.raw

```

179      290      295      300
181 Ala Gln Ala Leu Pro Gly Gln Met Leu Tyr Thr Gln Pro Pro Leu Lys
182 305      310      315
184 Phe Val Arg Ala Pro Met Val Val Gln Gln Pro Pro Val Gln Pro Gln
185      325      330      335
187 Val Gln Gln Gln Gln Thr Ala Val Gln Thr Ala Gln Ala Ala Gln Met
188      340      345      350
190 Val Ala Pro Gly Val Gln Val Ser Gln Ser Ser Leu Pro Met Leu Ser
191      355      360      365
193 Ser Pro Ser Pro Gly Gln Gln Val Gln Thr Pro Gln Ser Met Pro Pro
194      370      375      380
196 Pro Pro Gln Pro Ser Pro Gln Pro Gly Gln Pro Ser Ser Gln Pro Asn
197 385      390      395      400
199 Ser Asn Val Ser Ser Gly Pro Ala Pro Ser Pro Ser Ser Phe Leu Pro
200      405      410      415
202 Ser Pro Ser Pro Gln Pro Ser Gln Ser Pro Val Thr Ala Arg Thr Pro
203      420      425      430
205 Gln Asn Phe Ser Val Pro Ser Pro Gly Pro Leu Asn Thr Pro Val Asn
206      435      440      445
208 Pro Ser Ser Val Met Ser Pro Ala Gly Ser Ser Gln Ala Glu Glu Gln
209      450      455      460
211 Gln Tyr Leu Asp Lys Leu Lys Gln Leu Ser Lys Tyr Ile Glu Pro Leu
212 465      470      475      480
214 Arg Arg Met Ile Asn Lys Ile Asp Lys Asn Glu Asp Arg Lys Lys Asp
215      485      490      495
217 Leu Ser Lys Met Lys Ser Leu Leu Asp Ile Leu Thr Asp Pro Ser Lys
218      500      505      510
220 Arg Cys Pro Leu Lys Thr Leu Gln Lys Cys Glu Ile Ala Leu Glu Lys
221      515      520      525
223 Leu Lys Asn Asp Met Arg Cys Pro Leu Pro His Arg Pro Arg Cys His
224      530      535      540
226 Arg Pro Asn Ser Ser Thr Tyr Ala Ser Arg Ser Trp Met Pro Ser Trp
227 545      550      555      560
229 Pro Thr Ser Ala His Leu Ser Ser Thr Ile Pro Cys Thr Ala His Ser
230      565      570      575
232 Phe Gln Pro
236 <210> SEQ ID NO: 4
237 <211> LENGTH: 15
238 <212> TYPE: PRT
239 <213> ORGANISM: Homo sapiens
241 <400> SEQUENCE: 4
242 Arg Arg Met Ile Asn Lys Ile Asp Lys Asn Glu Asp Arg Lys Lys
243 1      5      10      15
246 <210> SEQ ID NO: 5
247 <211> LENGTH: 6
248 <212> TYPE: PRT
249 <213> ORGANISM: Homo sapiens
251 <400> SEQUENCE: 5
252 Gln Ser Ser Gln Ala Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,119

DATE: 10/09/2001  
TIME: 16:53:14

Input Set : A:\01100250.app

Output Set: N:\CRF3\10092001\I668119.raw

```

253 1 5
256 <210> SEQ ID NO: 6
257 <211> LENGTH: 107
258 <212> TYPE: PRT
259 <213> ORGANISM: Homo sapiens
261 <400> SEQUENCE: 6
262 Ala Leu Gln Gln Gln Gln Gln Gln Gln Phe Gln Gln Gln Gln Gln
263 1 5 10 15
265 Ala Ala Leu Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Phe Gln
266 20 25 30
268 Ala Gln Gln Ser Ala Met Gln Gln Gln Phe Gln Ala Val Val Gln Gln
269 35 40 45
271 Gln Gln Gln Leu Gln Gln Gln Gln Gln Gln Gln Gln His Leu Ile Lys
272 50 55 60
274 Leu Arg Arg Gln Asn Gln Gln Gln Ile Gln Gln Gln Gln Gln Leu
275 65 70 75 80
277 Gln Arg Ile Ala Gln Leu Gln Leu Gln Gln Gln Gln Gln Gln Gln
278 85 90 95
280 Gln Gln Gln Gln Gln Gln Gln Gln Ala Leu Glu
281 100 105
284 <210> SEQ ID NO: 7
285 <211> LENGTH: 25
286 <212> TYPE: DNA
287 <213> ORGANISM: Artificial Sequence
289 <220> FEATURE:
290 <223> OTHER INFORMATION: Description of Artificial Sequence: linker
291 sequence
293 <400> SEQUENCE: 7 25
294 gaattcagat ctcccggtc accgc
297 <210> SEQ ID NO: 8
298 <211> LENGTH: 20
299 <212> TYPE: DNA
300 <213> ORGANISM: Artificial Sequence
302 <220> FEATURE:
303 <223> OTHER INFORMATION: Description of Artificial Sequence: linker
304 sequence
306 <400> SEQUENCE: 8 20
307 tgaccggga gatcgattc

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/668,119

DATE: 10/09/2001  
TIME: 16:53:15

Input Set : A:\01100250.app

Output Set: N:\CRF3\10092001\I668119.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date